REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
21-05-2012	Final Report	1-Jul-2008 - 30-Jun-2012

4. TITLE AND SUBTITLE
Final Progress Summary for the Project "Development of a Novel Technology for Label Free DNA Sequencing"

5a. CONTRACT NUMBER
W911NF-08-1-0286

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
622622

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Peiji Zhao

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
North Carolina State University
Office of Contract and Grants
Leazar Hall Lower Level- MC
Raleigh, NC 27695 -7214

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
54616-EL.1

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
The accurate detection of biological material, such as antibodies, proteins or deoxyribonucleic acid(DNA), is currently receiving much attention for applications in molecular biology, medical sciences and military defense. Physical methods, such as infrared (IR) spectroscopy and THz spectroscopy, have been investigated for the detection of biological (bio) and organic materials. It is the goal of the proposed research to develop the theoretical foundation for the development of an optical/quasi-optical based technology for label free

15. SUBJECT TERMS
DNA sequencing, Codons, signatures

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		Peiji Zhao
					19b. TELEPHONE NUMBER: 919-515-3984

Report Title

Final Progress Summary for the Project
"Development of a Novel Technology for Label Free DNA Sequencing"

ABSTRACT

The accurate detection of biological material, such as antibodies, proteins or deoxyribonucleic acid(DNA), is currently receiving much attention for applications in molecular biology, medical sciences and military defense. Physical methods, such as infrared (IR) spectroscopy and THz spectroscopy, have been investigated for the detection of biological (bio) and organic materials. It is the goal of the proposed research to develop the theoretical foundation for the development of an optical/quasi-optical based technology for label free DNA sequencing or RNA expressions.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Number of Presentations: 2.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received	Paper

TOTAL:

Number of Manuscripts:

Books

Received	Paper

TOTAL:

Patents Submitted

Patents Awarded

Awards

Graduate Students

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Under Graduate students supported

NAME PERCENT_SUPPORTED

FTE Equivalent:

Total Number:

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 0.00

Names of Personnel receiving masters degrees

NAME

Total Number:

Names of personnel receiving PHDs

NAME

Total Number:

Names of other research staff

NAME PERCENT_SUPPORTED

FTE Equivalent:

Total Number:

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

Merit of the Research:

1) The suggested technology is the first high throughput technology based on the IR/THz radiation and electronics.

2) The capability of the sequencing (or the detection) of DNA samples with ultra-low concentration of DNA molecules. As we know that the number of DNA molecules in a sample is very important for most of the DNA sequencing techniques. If the number of DNA molecules is too small, the number of the molecules in the sample has to be amplified by using PCR technique. In the proposed technique, because of the structure of the device (break junction structure), there is no need for the amplification of the number of DNA molecules in the sample in the preparation of the sample for DNA sequencing.

3) The very high accuracy and the capability of positioning the base pair mismatches and lesions in DNA molecules. The very high accuracy and capability of positioning the base pair mismatches and lesions in DNA molecules is another important characteristics of the proposed technology. The proposed technique extracts the sequence information by investigating the relationship between the radiation modulated current and the sequence of the bases in the molecules. Here, the vibrational characteristics of the molecules greatly influence the currents through the DNA molecules. Furthermore, the vibrational characteristics of the molecules are determined by the sequence of the bases and the conformation of the molecules. Hence, the current through the
molecule will be greatly influenced by the vibrational characteristics associated with the sequence of the bases in the molecule. The current through the molecule should carry the accurate sequence information of the DNA molecule and reflects the exact positions of the mismatches and lesions.

The reusability of the device. The reusability of the device used in the suggested technique is a great advantage comparing to other DNA sequencing techniques. In current label free sequencing techniques, such as the techniques based on carbon nanotubes and DNA microarrays, the devices have to be discarded once they are used for sequencing of DNA or RNA expression. In the suggested technique, the molecule bound on the electrodes can be flashed away after the device is used for sequencing of DNA or RNA expression. Hence, the suggested technique can greatly reduce the cost of the DNA sequencing, which is one of the central concern or requirement for current DNA sequencing techniques.

Please see the attached report for details on the accomplishments.

Technology Transfer

Final Progress Summary for the Project
"Development of a Novel Technology for Label Free DNA Sequencing"

Peiji Zhao and Kurt Becker – Polytechnic Institute of New York University

Overall Project Goal:

The accurate detection of biological material, such as antibodies, proteins or deoxyribonucleic acid (DNA), is currently receiving much attention for applications in molecular biology, medical sciences and military defense. Physical methods, such as infrared (IR) spectroscopy and THz spectroscopy, have been investigated for the detection of biological (bio) and organic materials. Although IR spectroscopy is useful for probing the secondary structure of DNA/RNA, it cannot provide base pair sequence information, which is necessary for the accurate detection and identification of bio-materials. Here, the statistical effects associated with interacting molecules tend to smear out all sequence information of the individual DNA/RNA targets. Therefore, the current optical and quasi-optical techniques are not adequate for collecting detailed sequence information from DNA/RNA molecules***. DNA sequencing and RNA expression in terms of optical or quasi-optical techniques is a long-term unsolved important scientific and engineering issue having widespread application both in medicine and biology and in homeland defense with huge political and economic impacts.*** It is the goal of the proposed research to develop the theoretical foundation for the development of an optical/quasi-optical based technology for label free DNA sequencing or RNA expressions.

Research Goals:

The goals of the research project involve the development of the fundamental relationship between the current through a DNA molecule and the sequence of the nucleotides inside the molecule. Specifically, the proposed research aims to the understanding to the following important physical issues which renders the relationship between the current and the sequence of the nucleotides.

- Dynamical Characteristics of Codons for RNA expressions and DNA Sequencing, Energy Structures of Codons
- Interactions Among Codons and Relationship Between Sequence of the Bases and the Vibrational Modes
- Molecular Dynamics Based Low Frequency Oscillations in DNA Bound on Surface of Electrodes and Codon Interactions
- Motion of Holes in DNA for DNA Sequencing
- Relationship between the radiation-modulated current through a DNA molecule and the sequence of the nucleotides inside the molecule

Impact of the Research:

If the proposed research is successful, the proposed research would have several important impacts to the bio-medical engineering, biophysics, and the technology for the detection and discrimination of bio-agents. Specifically, the expected impacts involve those in economy and politics. To this connection, the most important impact of the project is the relationship between the current through a DNA molecule and the sequence of the nucleotides inside the molecule. As we have discussed previously, present optical spectral based technologies cannot provide information relation between the optical spectra and the sequence of nucleotides inside the molecule. DNA sequencing and RNA expression in terms of optical or quasi-optical techniques is a long-term unsolved important scientific and engineering issue. The proposed research aims to resolve this issue for next generation DNA sequencing technology. Here, the swiftness and the accuracy of optical/quasi-optical based technologies in the detection and discrimination of bio-agents is an important advantage comparing with other technologies. Hence, considering the importance of DNA sequencing technologies in modern medicine and homeland defense, the proposed research of the technology will have widespread application both in medicine and biology and in homeland defense with huge political and economic impacts.

Technology Transfers & Research Collaborations:

This research is done in collaboration with Dr. Ena Wang and Dr. Francesco Marincola at NIH. This collaboration ensures the correctness in biology.

Merit of the Research:

- *The suggested technology is the first high throughput technology based on the IR/THz radiation and electronics*.
- *The capability of the sequencing (or the detection) of DNA samples with ultra-low concentration of DNA molecules*. As we know that the number of DNA molecules in a sample is very important for most of the DNA sequencing techniques. If the number of DNA molecules is too small, the number of the molecules in the sample has to be amplified by using PCR technique. In the proposed technique, because of the structure of the device (break junction structure), there is no need for the amplification of the number of DNA molecules in the sample in the preparation of the sample for DNA sequencing.
- *The very high accuracy and the capability of positioning the base pair mismatches and lesions in DNA molecules.* The very high accuracy and capability of positioning the base pair mismatches and lesions in DNA molecules is another important characteristics of the proposed technology. The proposed technique extracts the sequence information by investigating the relationship between the radiation modulated current and the sequence of the bases in the molecules. Here, the vibrational characteristics of the molecules greatly influence the currents through the DNA molecules. Furthermore, the vibrational characteristics of the molecules are determined by the sequence of the bases and the conformation of the molecules. Hence, the current through the molecule will be greatly influenced by the vibrational characteristics associated with the sequence of the bases in the molecule. The current through the molecule should carry the accurate sequence information of the DNA molecule and reflects the exact positions of the mismatches and lesions.
- *The reusability of the device*. The reusability of the device used in the suggested technique is a great advantage comparing to other DNA sequencing techniques. In current label free sequencing techniques, such as the techniques based on carbon nanotubes and DNA microarrays, the devices have to be discarded once they are used for sequencing of DNA or RNA expression. In the suggested technique, the molecule bound on the electrodes can be flashed away after the device is used for sequencing of DNA or RNA expression. Hence, the suggested technique can greatly reduce the cost of the DNA sequencing, which is one of the central concern or requirement for current DNA sequencing techniques.

Summary of Achievements: Tasks completed in the period include,

- **Achievement #1: Characteristics of Charge Transfer Through the Exon 2 in Tumor Suppresser Gene TP53**

  The conductivity of DNA materials has been studied for many years for the importance of the research subject in biology and in molecular electronics. While much attention has been focused on the study of the conductive characteristics of artificial DNA molecules, such as poly(dG)-

poly(dC), etc., less attention has been focused on the charge transfer inside a human genetic molecule. Here, it should be noted that artificially fabricated DNA molecules, such as poly(dG)-poly(dC), etc., are not genetic molecules because they don't carry genetic information, although they may have important implementations in nano-electronics and bio-physics for the explanation of the mechanism of the conduction of the molecules. The study of conductive characteristics of larger genetic molecules is of great importance not only to the explanation of some biological processes such as the natural healing of genetic mutation in a gene but also to the development of the novel ultrahigh throughput DNA sequencing technique we suggested before.

Biological study shows that the genetic information is stored in the exons of a genomic molecule. An exon is composed of a number of DNA codons aligned in a linear form for the storage of genetic information. Here, the DNA codons are actually triplets of the nucleotides, such as GCT, ACT, etc. The codons store information for protein syntheses. In practical biology processes, it is the change of the nucleotide sequence of the corresponding codons that causes the mutation of the exon. It is reasonable to assume that the change in DNA base sequence is communicated through electric signals inside a biological object. Hence, the study of the characteristics of the electric conduction of exons is the key in getting the insight into the conductive characteristics of genetic molecules and charge transfer process inside the molecules. This knowledge is of greatly important to the explanation of the natural healing of genetic mutation. *To this connection, the relation between the mutation sites and the electric characteristics of the exons will provide protocols for the development of new generation DNA sequencer.* Indeed, our theoretical calculation results show clear changes of the transmission coefficients of the exon 2 of the tumor suppresser gene P53. Furthermore, the changes of the mutation are closely related to the positions of the mutation sites in the molecule. This is an important result to electric-signal-based DNA sequencing techniques.

In this study, by using tight-binding model, a well accepted model for the study of charge transfer in DNA, we systematically study the conductive characteristics of the exon 2 of the tumor suppresser gene TP53. The relationship between the sequence of bases in the molecules and the electric characteristics of the molecules was studied. The main research results are listed below.

1) *The mutations of the exon are always associated with the increase of the transmission coefficient of the charge through the exon*, *with one exception when mutation site is at codon* 22 (CTA → CTG). Specifically, when a mutation occurs such that a base with a low on-site energy is replaced by a base with a larger on-site energy, the transmission coefficient of the charge increases, vice versa.

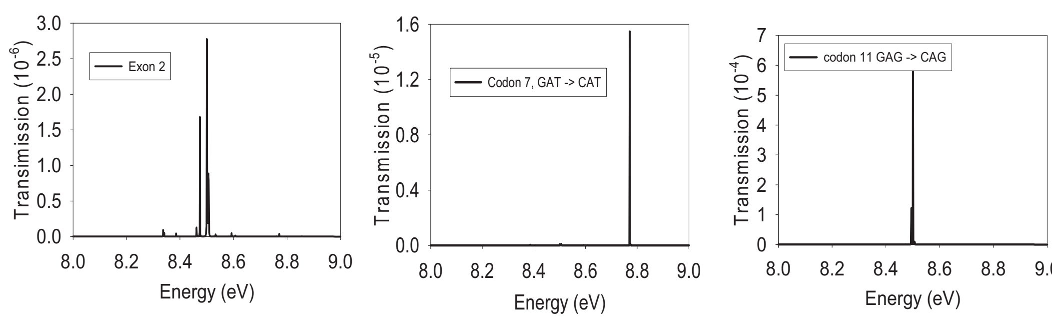

Figure 1 (a) Transmission coefficient of exon 2 of tumor suppress gene TP53; (b) transmission coefficient of the exon 2 carrying mutation at codon 7; (c) transmission coefficient of exon 2 bearing mutation at codon 11.

Final Progress Summary for the Project
"Development of a Novel Technology for Label Free DNA Sequencing"

2) *The 10^x power rules of the transmission coefficient of the mutated exon 2*. The increase in the transmission coefficient and the type of mutations has shown a 10^x power rules. The relationship can be illustrated in terms of the site index of the bases in a codon table. If we let i be the first letter of the codon, j be the second letters, and k the third letter of the codons. Then, the index i(j, k) = 1, 2, 3, and 4 are associated with T, C, A, G. To this connection, the magnitude of the increase is a factor of $10^{|i_2-i_1|}$ higher than that of the unmutated exon if the mutation is related to the change of the first letter of a codon. Similarly, the increase is a factor of $10^{|J_2-J_1|}$higher than that of the unmutated exon if the mutation is associated to the variation of the second letter of a codon. Furthermore, the change of the transmission coefficients is a factor of $10^{|k_2-k_1|-1}$ if the third letter is the mutation site. There is an exception to this rule for the mutation GAT → CAT. Here, the reduction in the magnitude of the transmission coefficient may be due to the larger value of the on-site energies of A and T. Further investigation is needed to clear the difference between the rule and the calculated result.

3) *An assumption on the mechanism of the natural healing of damaged DNA*. The mutations also create(or close) some marker channels for the transfer of the charge in a mutated exon. As we have discussed, the electrical signals are important to the natural healing of a damaged DNA. Here, if the transmission coefficient of a mutated exon does not change or change very small comparing to the value of the transmission coefficient of the original exon, the system can not detected the change of the base sequence. Therefore, the mutated codon cannot be repaired by the corresponding repairing mechanism. From this point of view, we can assume that the cells with the mutations GAG → GAA or CTA → CTG may lead to tumor and cannot be repaired by the base excision repair enzymes because the enzymes are not able to feel the mutations in the molecule. For other mutations, the resulting transmission coefficients either are large comparing to the original value of the coefficient or have new spectra. Therefore, the cell systems may detected the change in the coefficient. Sometime, the mutations still exist and the damaged DNA are not be repaired. In fact, the process is related to the complicated biological processes and cannot be solely explained by the tight binding approximation. In real processes, the vibrations of the lattice and the environmental variables are crucial factors influencing the charge transfer inside a DNA molecule. A fine model of charge transfer in a DNA molecule is needed for the understanding to the basic feature of the conduction of the charges in the exon. The conduction channels for each mutated exon 2 are listed in the table below.

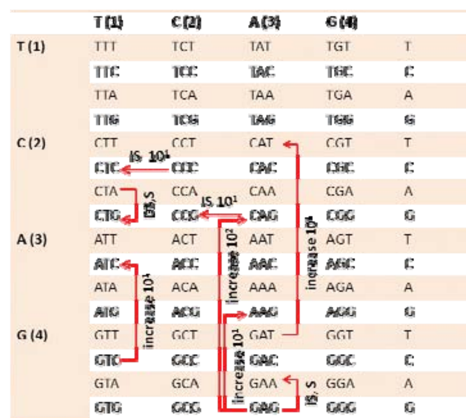

Figure 2 Graphic illustration of the relation of the increase of the transmission coefficients and the mutation sites.

Final Progress Summary for the Project
"Development of a Novel Technology for Label Free DNA Sequencing"

Energy(eV)	Native Exon 2	codon 2 GAG->GAA	codon 5 CAG->CCG	codon 7 GAT->CAT	codon 10 GTA->ATC	codon 11 GAG-> CAG	codon 11 GAG-> AAG	codon 12 CCT->CTC	codon 18 ACA -> ACT	codon 22 CTA -> CTG
8.2163						0.2584				
8.2557		1.22E-01								
8.3333									10.025	
8.3371		0.3343								
8.3397							0.4012		0.7231	
									0.4195	
8.3409		2.07E-01								
8.3893								5.7694		
8.3855		0.1057								
8.4618	0.1273	0.177	7.3553						0.4958	
8.463			0.1279						0.1976	
8.4643					0.1189					
8.4656								18.336		
8.4669			0.1227							
8.4681			0.4912							
8.4694			9.7665							
8.4707			0.2089				0.3142			
8.4719							0.6915			
8.4732						0.7166	0.1334		0.1309	
8.4745	1.6811	2.3118					0.1037	0.5699	2.999	0.9514
8.4758					1.9965		0.1509	0.1273		
8.477					0.2473		0.5174			
8.4783							20.077			
8.4796							0.234			
8.4872			0.2303							
8.4885			2.2702							
8.4898			1.2717							
8.491			0.2396			0.1236				
8.4923			0.1129			0.4177				
8.4936						3.0753				
8.4948						122.05				0.1221
8.4961						4.7776				1.317
8.4974			0.1021		0.1048	3.08	0.1032			
8.4987	0.1428	0.1534	0.2272		0.3764	4.3335	0.3466			
8.4999	2.7791	2.9563	2.6213		6.7043	15.576	4.9775	1.4673		
8.5012	1.0984	1.1573	0.5907		3.013	650.66	3.2607	0.6177		
8.5025	0.2673	0.279			0.6978	9.8454	0.6638	0.1408		
8.5037	0.1918	0.1983			0.488	3.8965	0.4409			
8.505	0.2571	0.2635			0.6396	3.6114	0.5572	0.1228		
8.5063		0.9008		0.1022	2.1546	9.366	1.8256	0.4033		0.1278
8.5076		0.7849			1.9427	6.7641	1.567	0.3465		0.1355
8.5088					0.2112	0.6092	0.1666			
8.5101						0.1419				
8.533									0.9712	
8.5355					1.5705					
8.5457								1.9244		
8.5648									0.6501	
8.6068									5.705	
8.6323									0.325	
8.6411								0.1651		
8.7708				15.483						
8.7747									8.4937	
8.8637									10.722	

Achievement #2: First Principle Study of the Terahertz and Far-Infrared Spectral Signatures in DNA Bonded to Silicon Nanodots

Previously, we have studied the electric characteristics of an organic molecule bonded on the surface of silicon. This research topic is of importance to the proposed research on the development of the novel DNA sequencer. Our results illustrate the basic rules for the design of the electrostatic characteristics of the contact between the molecules and the silicon surfaces as shown in Figure 7. The dashed line in the figure is the reference line for silicon surface. For noncyclical attachment, the tether atoms above this line have greater electronegativities than that of silicon atoms. They will create barriers between the linker molecules and the surface, which form a Schottky-like contact for electrons. Below this line, the electronegativities of the tether atoms are slightly less than that of silicon atoms. These atoms create charge traps instead of barriers between the linker molecule and the surface for electrons, which form an Ohmic-like contact. An Ohmic -like contact can also be created by using a cyclical attachment where the electronegativity of the stand-by atom (the nitrogen atom) is greater than those of the carbon and silicon atoms. These results suggest a way for the design of a stable and electrically effective contact between a molecule and the surface of silicon. The electrostatic characteristics of the contact can be modulated by adjusting the electronnegativity of the stand-by atom in the side branch of a cyclical attachment structure. This design variation is a noteworthy and one that will have important utility in practical applications. Although the study shows the fundamental electric characteristics of the contacts and the design methodology, the atomic geometric structures of the contact are not considered in the calculations. Furthermore, the IR/THz spectral characteristics of the contact is not studied although the spectra is of important to the development of the new IR/THz wave modulated current for sequencing. In the current study of the issue of organic molecule bonded on silicon electrodes, we studied the infrared and THz spectral signatures of DNA molecules bonded on the surface of a nanodot that is the model of silicon electrodes. The main simulation results are listed below.

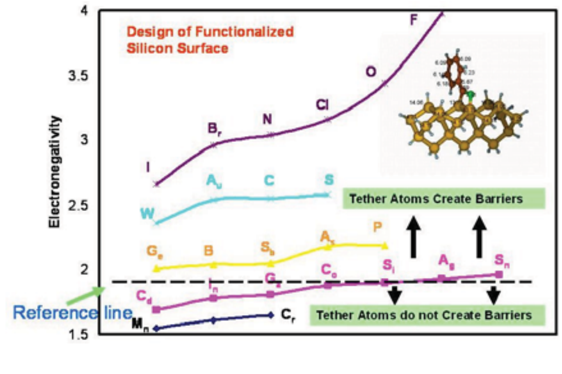

Figure 3 Design rule for the electric characteristics of the tether atom and the silicon surface.

1) The DNA fragments (dG or dGG) bonded to silicon (111) nanodots do not cause large deformations in silicon nanodots, which makes them suitable for traditional device applications from the technology perspective.

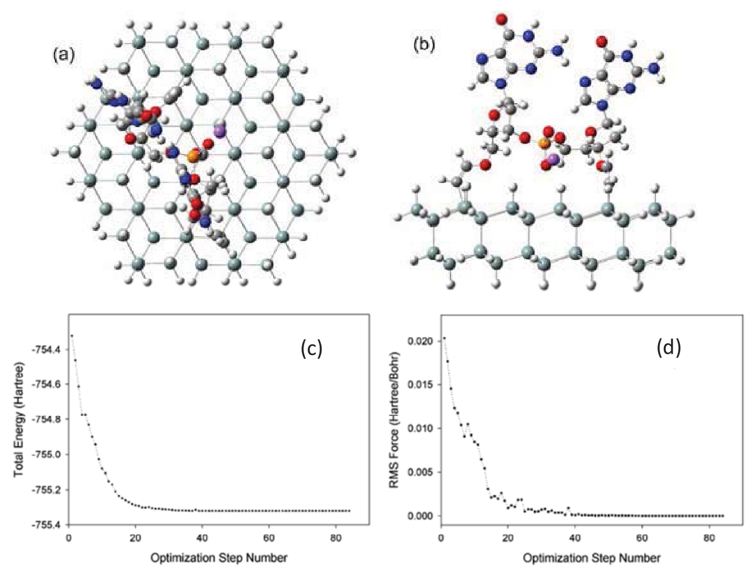

Figure 4 Stability of DNA hybrid system

2) The optically active modes lying within the terahertz spectrum typically arise out of joint coupling between the DNA's vibrational behavior and the dynamics within the nanodot substrate. For many normal modes, this coupling is predicted to be strong and can interfere with attempts to spectrally characterize DNA target molecules.

Final Progress Summary for the Project
"Development of a Novel Technology for Label Free DNA Sequencing"

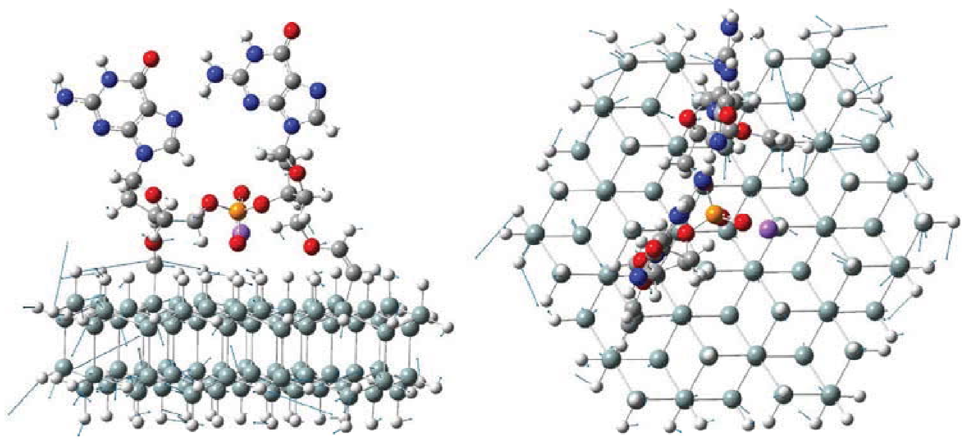

Figure 5 Optically active mode at ~1 THz (132.1 cm $^{-1}$) with very strong coupling between dGG and Si. Views from side and above are shown. Thin lines show the directions of atomic displacements

3) For the single strand dGG, the existence of some weakly-coupled, normal-modes that are intrinsic to the structure of the dGG molecule. In particular, the dominant absorption line below 6 THz (200 cm^{-1}) is predicted to strongly involve the DNA and effects of sodium, but it is only weakly influenced by the silicon nanodot vibrations and associated edge effects.

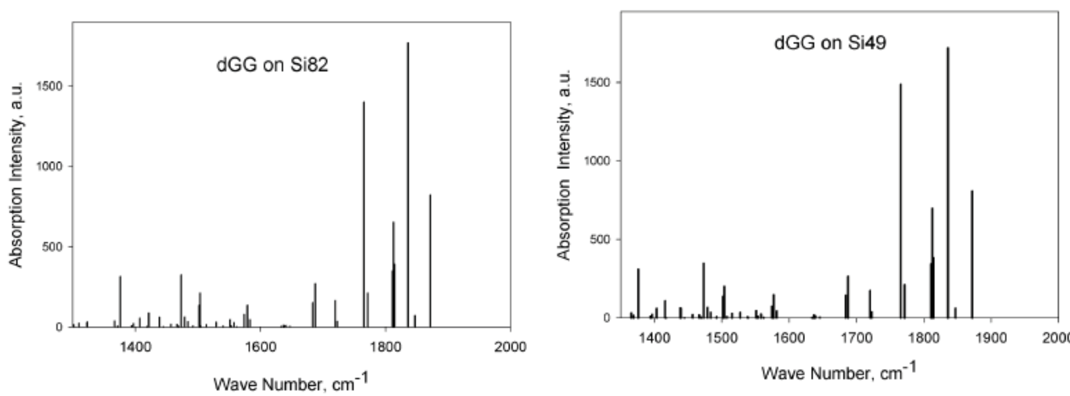

Figure 6 Absorption spectra in the far-IR. Simulation results for Si82 nanodot ("dGG on Si82") and Si49 nanodot ("dGG on Si49").

4) Spectral signatures strongly depend on DNA chain length at THz frequencies. The DNA maintains some unique THz signatures that are independent of the Silicon nanodot geometry. Hence, one can fairly speculate that THz spectral signatures useful for fingerprinting DNA molecules should also exist for the case of longer single-strand DNA fragments and/or DNA molecules of differing base composition when bonded into DNA/nanodot systems. Therefore, these studies have shown that THz spectral signatures exist in hybrid DNA/nanodot systems with the general characteristics needed for the successful implementation of ultrasensitive nanoscale detectors.

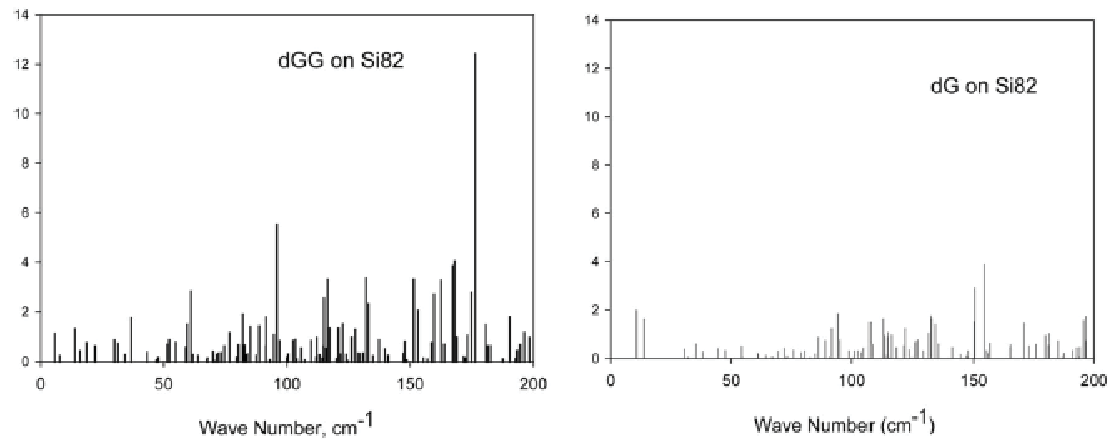

Figure 7 Influence of the length of DNA molecules on the vibrational spectra of the molecules.

5) When DNA fragments are bonded to silicon nanoprobes and subjected to active THz illumination, the unique THz-frequency phonon vibrations will be generated that could be sensed by direct electron current measurements.

➢ Achievement #3: Vibrational Characteristics of Genetic Codons: 5'-GGX-3' and 5' -XGG-3' (X = A, C, and T)

The vibrational characteristics of DNA molecules is critical to the study of the relationship between the IR/THz wave modulated current through the molecule and the sequence of the nucleotides in the molecule. To this connection, we have theoretically investigated the vibrational characteristics of genetic codons for the significance of the codons in the storage of bio-information in DNA molecules. This study reveals a number of phenomenological trends that have scientific relevance to spectroscopic characterization of DNA molecules. The main results are listed below.

1) The spectra of the six DNA codons (i.e., at the sequence forms 5' -GGX-3' and 5' -XGG-3' (X = A, C, and T) in gas phase in the near infrared regime can be classified into five semi-distinct spectral sub-regimes, namely: two end regions and one region that separates the two sequence identifier regions with features that arise primarily from coupling between the GG bases and other bases (T, C, and A).
2) The spectra in the high-frequency end-regions arise primarily from vibrational modes associated with the O-H bonds at the 5'end or 3' end of the codons and the spectra in the low-frequency end-regions are formed primarily by the vibrational modes of the backbone.
3) In the lower-frequency side, the sequence-identifier region is composed of the C-H bond stretch vibrations in the planes of the corresponding DNA bases, and in the higher-frequency side, sequence-identifier region is composed of the N-H bond stretch vibrations in the planes of the corresponding DNA bases. In addition, the sequence-identifier dividing region almost exclusively contains vibrational modes due to coupling between the G nucleotides and other bases.
4) All the vibrational modes in sequence identifier regions are localized at the corresponding DNA bases and exhibit a definable dependence on the sequence form of the codons under study.

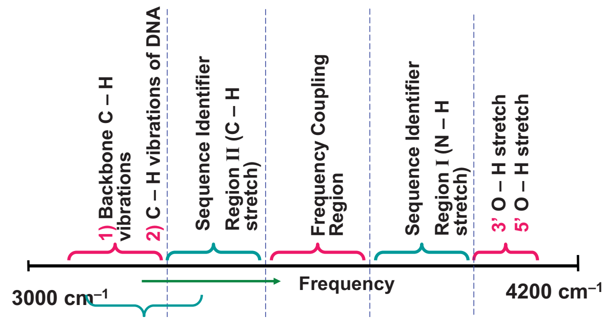

Figure 8 Graphic illustration of the specral structures of the genetic codons in near infrared region.

- **Achievement #4: Chaotic and Ultrahigh Frequency Rabi Oscillation in a Coupled-Double-Quantum-Dot Semiconductor System**

The mechanism of the charge transfer inside a DNA molecule under the influence of THz radiation is crucially important to the study of the relationship between the IR/THz modulated current and the sequence of the nucleotides. As well accepted, the G DNA plays a role of quantum well and T DNA acts as a potential barrier in the process of charge transfer inside a DNA molecule. To this connection, in terms of density matrix theory, we have studied the Rabi oscillation of the charge transfer in a coupled-double-quantum-dot system under a high THz field. The research results are not only suitable to semiconductor system but also to DNA molecular system. For example, a DNA molecule with sequence of ATGGTGGTA is a counterpart of the semiconductor coupled quantum dot system. The main results are listed below.

1) The probability oscillation exhibits chaotic features as well as the characteristics of collapse and revivals of the oscillations. Here, the nonlinear characteristic is due to the multi-photon excitations of the electron.

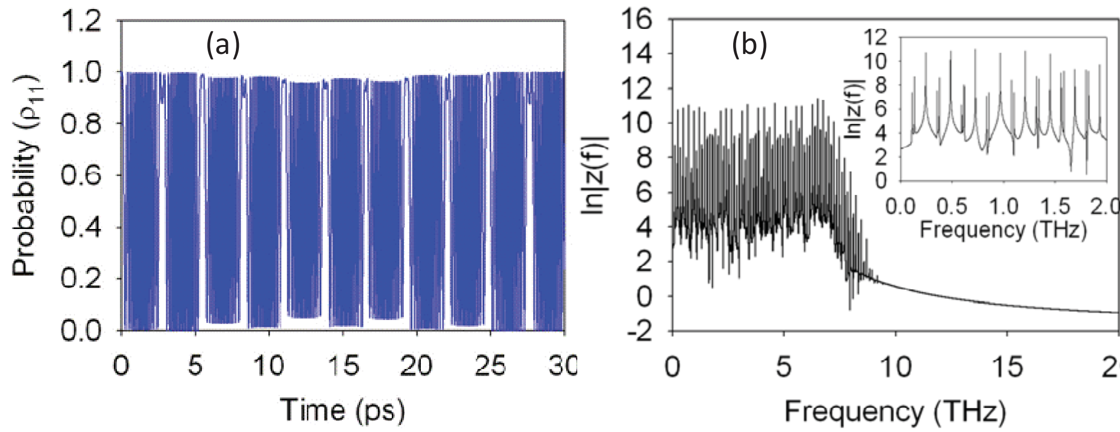

Figure 9 (a) The step-like feature of the ultrahigh frequency oscillation is one of the indication of chaotic oscillation; (b) The power spectrum of the probability is one of the evidence of the creation of chaos.

2) The coherence time of the electron may be greatly extended through the collapse and revival mechanism of the oscillations. The chaos-enhanced coherence may have important application in quantum information processing.

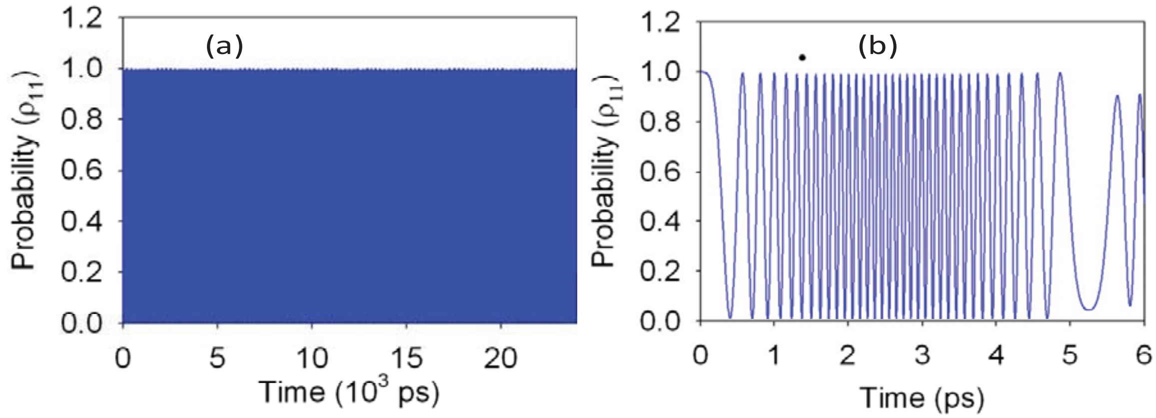

Figure 10 (a) Low resolution of the oscillation; (b) higher resolution of the oscillation.

3) The creation of the ultrahigh frequency oscillation of the electronic probability in each probability step. This characteristics of the oscillation may have important application in ultrafast quantum information processing.

- **Achievement #5: Investigation of the Intrinsic Conduction Mechanism of Charge in DNA Molecules.**

In our current research, we found that the number of the conductive channels of the electrons in a DNA molecule and the magnetite of the transmission coefficients is greatly related to the length of molecule or the number of bases in the molecule. More interestingly, the conduction characteristics of the charge verse the length of the molecule is not a monotonic function. The figure below shows the relationship between the transmission coefficients of the charge and the number of bases clearly. This result is very important to information transfer in DNA/RNA and in bio-electronics. It should be noted that the result is obtained in terms of the tight binding theory which ignores the detail of the interaction between the nucleotides and that between the bases and the environment. Currently, we are

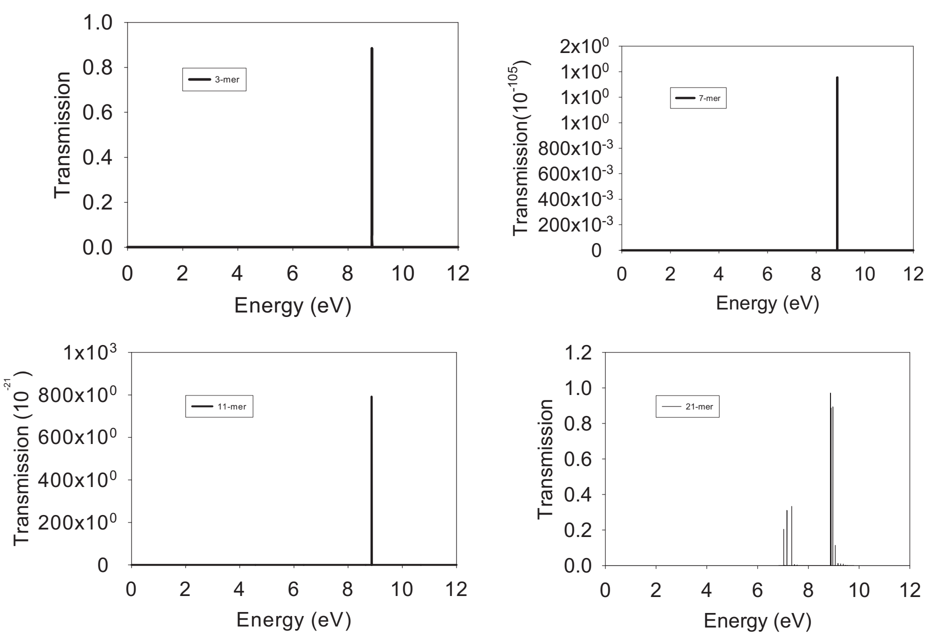

Figure 11 The relation between the number of conduction channels and the length of the DNA molecules.

studying this issue in terms of density functional theory and Green's function theory for transport of charges inside a DNA molecule. This study will reveal the answer to this issue and predict relationship between the conductivity and the length of the molecule and the environment.

Achievement #6: Study of the Physical Mechanism of DNA Damage Recognition

DNA damage repair refers to a collection of biological processes by which a cell identifies and corrects damage to the DNA molecules that encode its genome. Biologically, the damaged DNA must be repaired to prevent loss or incorrect transmission of genetic information, as errors can cause developmental abnormalities and tumorigenesis. There are three major DNA repairing mechanisms: base excision, nucleotide excision and mismatch repair. Generally, a DNA repair process involves the identification of the damaged bases and the replacement of the damaged bases. Here, the enzymes known as DNA glycosylases remove damaged bases by literally cutting them out of the DNA strand through cleavage of the covalent bonds between the bases and the sugar-phosphate backbone. The resulting gap is then filled by a specialized repair polymerase and sealed by ligase. Currently, most of the research focus on the second step of the biological process for DNA repair. There is no serious research focus on how does a damaged DNA base be found in the genome although researchers knew that the enzymes can figure out which bases. However, there is no answer to the question: how a enzyme molecule can find the site of the damaged bases in the genome. In this study, we will present a mechanism for the explanation of the question: how does the surveillance system of a cell find the site of the damage? Here, a charge transfer inside a DNA molecule plays a vital role in the surveillance system of a cell for locating the site of the damaged nucleotide. Specifically, the difference of the charge distributions in a normal gene and a damaged gene would change the electric field around the molecule. To this connection, the molecule of excision enzymes can feel the difference of the difference of the charge distribution through electric field interaction. Once the difference in charge distribution is detected, a repair process will be started. Our numerical calculation results as shown in Figs.(12) do shown the difference of the charge distribution in resonant transfer of an excess charge through the exon 2 of tumor suppresser gene TP35. Biologically, this study is important to the study of biological aging process as well as prevention of cancers. This study is also important to the development of the novel technology for label free DNA sequencing for the detection of the sites of DNA base mutation or lesions for the damage will change the interaction between the molecule and the electrode.

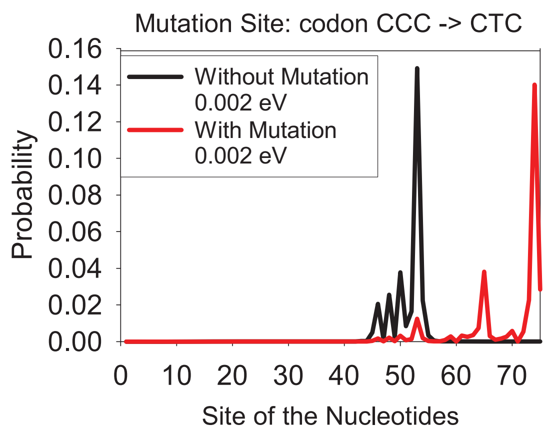

Figure 12 Mutation at the codon 12 (CCC ->CTC) changes the distribution of the electron density inside the exon 2 of the tumor suppressor gene TP35 when the energy of the injected electron is 2 meV which is in the range of thermal excitation..